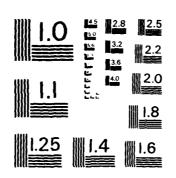
AD-A13.1 986 UNCLASSIFIED	USE IN THE TH	HREE(U) AII D'ENTREMONT E	MAGERY ANALYS R FORCE GEOPH T AL. 28 DEC 1	YSICS LAB H	ANSCOM	1	
END DATE FRAMED *9 -83 DTIC							



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS - 1963 - A

. . . . . .

AFGL-TR-82-0397 ENVIRONMENTAL RESEARCH PAPERS, NO. 817





Evaluation of Automated Imagery Analysis Algorithms for Use in the Three-Dimensional Nephanalysis Model at AFGWC

ROBERT P. d'ENTREMONT RUPERT S. HAWKINS JAMES T. BUNTING

28 December 1982



Approved for public release; distribution unlimited.

METEOROLOGY DIVISION

PROJECT 6670

AIR FORCE GEOPHYSICS LABORATORY

HANSCOM AFB, MASSACHUSETTS 01731

AIR FORCE SYSTEMS COMMAND, USAF



83 08 25 068

UTC FILE COPY

This report has been reviewed by the ESD Public Affairs Office (PA) and is releasable to the National Technical Information Service (NTIS).

This technical report has been reviewed and is approved for publication.

Oka J Stain, Jr.

DR. ALVA T. STAIR, Jr.

Chief Scientist

Qualified requestors may obtain additional copies from the Defense Technical Information Center. All others should apply to the National Technical Information Service.

# Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Date Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
AFGL-TR-82-0397	3. RECIPIENT'S CATALOG HUMBER
TITLE (and Subtitie) EVALUATION OF AUTOMATED IMAGERY ANALYSIS ALGORITHMS FOR USE IN THE	Scientific. Interim.
THREE-DIMENSIONAL NEPHANALYSIS MODEL AT AFGWC	6. PERFORMING ORG. REPORT NUMBER ERP No. 817
Robert P. d'Entremont Rupert S. Hawkins James T. Bunting	8. CONTRACT OR GRANT NUMBER(s)
Air Force Geophysics Laboratory (LYS) Hanscom AFB	10 PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT HUMBERS
Massachusetts 01731  CONTROL NO OFFICE NAME AND ADDRESS Air Force Geophysics Laboratory (LYS)	66701701 12 REPORT DATE 28 December 1982
Hanscom AFB Massachusetts 01731	13 NUMBER OF PAGES
4. MONITORING AGENCY NAME & ADDRESS: If different from Controlling Office)	15 SECURITY CLASS of this report:
	Unclassified
	15% DECLASSIFICATION DOWNGRADING
Windows Control of the Control of th	
The same of the sa	
Satellite cloud analysis Infrared i Automated cloud analysis Image pro Cloud layer analysis 3D Nepha	ncessing

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE When Data Entered

## SECURITY CLASSIFICATION OF THIS PAGE(When Date Entered

### 20. Abstract (continued)

percent cloud (earth) coverage and cloud heights. Both algorithms were found to perform well under some meteorological circumstances, but not so well under others. In general, however, it was found that CLUSTER performs better. The CLUSTER algorithm found clear/cloud boundaries, numbers of cloud layers, and fractional cloud covers that are closer to estimates made by analysts looking at corresponding IR and visible images. The MODAL algorithm found more cloud layers than were observed, and tended to overanalyze the transition regions between cloud layers.

Despite the overall advantages of CLUSTER, there is considerable room for improvement and algorithm development leading to improved cloud analysis. For instance, CLUSTER finds too many layers of cirrus clouds due to differing transparencies and sub-pixel sizes of cirrus cloud elements. It is possible to improve the cirrus layering problem by tuning the algorithm to be less sensitive to smaller thermal gradients at cold temperatures. CLUSTER can be tuned for other cloud types and levels

as well.

The AFGWC has recognized the limitations of MODAL and the complications in its computer code. Subsequently, a substantially revised version of the 3DNEPH has been written, and is known as the Real-Time Nephanalysis (RTNEPH). With the continued cooperation of the AFGWC, we plan to evaluate this new program's image analysis algorithm and recommend changes and improvements for it, drawing on our experience with MODAL and CLUSTER.

Accomples to r

NTIS tease
DITO TOS

Uncase to the second of the second

Unclassified

RITY CLASSIFICATION OF THIS PAGE When Data Entered

		Contents
1.	INTRODUCTION	5
2.	DETAILS OF THE ALGORITHMS	6
	<ul><li>2.1 Introduction and Details of MODAL</li><li>2.2 The CLUSTER Algorithm</li></ul>	7 10
3.	SUMMARIES OF THE DATA SET	12
	3.1 Categories of Clear and Cloudy Regions	14
	<ul> <li>3.2 Evaluation of the Two Algorithms in Terms of the Seven Basic Categories</li> <li>3.3 Frequencies of Clusters and Modes</li> <li>3.4 Preliminary Cloud Cover Evaluation</li> </ul>	17 31 34
4.	CONCLUSIONS	36
RE	FERENCES	30
		Illustrations
1.	MODAL Histogram Analysis	7
2.	Histogram Data Before and After Clustering With a Summary Below of the Four Clusters	13
3.	MODAL and CLUSTER Analysis for Case 302	15

		Illustrations
4	MODAL and CLUSTED Analysis for Case 167	18
4.	MODAL and CLUSTER Analysis for Case 167	
5.	MODAL and CLUSTER Analysis for Case 348	19
6.	MODAL and CLUSTER Analysis for Case 60	20
7.	MODAL and CLUSTER Analysis for Case 312	22
8.	MODAL and CLUSTER Analysis for Case 342	23
9.	MODAL and CLUSTER Analysis for Case 15	24
10.	MODAL and CLUSTER Analysis for Case 16	25
11.	MODAL and CLUSTER Analysis for Case 217	26
12.	MODAL and CLUSTER Analysis for Case 62	27
13.	MODAL and CLUSTER Analysis for Case 209	28
l4.	MODAL and CLUSTER Analysis for Case 301	29
15.	MODAL and CLUSTER Analysis for Case 199	30
16.	MODAL and CLUSTER Analysis for Case 349	32
17.	Percent Distribution of the Numbers of Clusters/Modes for Entirely Clear Cases and Cloudy Cases With Only One Category of Cloud Per Case	33
		Tables

		lable
1.	Options for Representative Grayshades of a Mode	8
2.	Corresponding Interval Ranges for the 6-Bit DMSP Data Used in This Study	12
3.	Composition of the Cloud Imagery Sample Set	16
4.	Cloud Cover Within 10 Percent of Observed	35
5.	Cloud Cover Closest to Observed	35

Evaluation of Automated Imagery Analysis Algorithms for Use in the Three-Dimensional Nephanalysis Model at AFGWC

### 1. INTRODUCTION

An automated satellite cloud analysis system has been in use for over a decade at the Air Force Global Weather Central at Offutt AFB, Nebraska. This system is known as the 3DNEPH (Nephanalysis) Model and has been described by Coburn<sup>1</sup> and Fye. <sup>2</sup> It is a very comprehensive system that merges satellite imagery information and conventional meteorological data to produce global cloud analyses eight times a day. The infrared processor is an important component of this system, since the infrared image is the sole source of satellite information at night. Frequently this is the case during the day as well.

The major component of the infrared processor is an algorithm called MODAL that was designed to separate major regions or layers in a  $25 \times 25$  n mi grid box which consists of an array of  $8 \times 8$  infrared values having a nominal 3 n mi resolution. The MODAL algorithm is a histogram separation procedure that gives from one to four distinct regions or layers for each grid box. Our

Received for publication 22 December 1982

- Coburn, V. R. (1971) Improved Three-Dimensional Nephanalysis, AFGWC Technical Memorandum 71-2.
- 2. Fye, F. K. (1973) The AFGWC Automated Cloud Analysis Model, AFGWC Technical Memorandum 78-002.

objective is to try to understand this algorithm better. What are its weaknesses? Where are its strengths?

An algorithm developed by Hawkins<sup>3, 4</sup> to analyze satellite data operates on infrared images, and can be configured to give much the same type of output as MODAL. This algorithm, called CLUSTER, separates the infrared images into mutually exclusive regions. All points are assigned to a region, whereas in MODAL some points may be left unassigned. In CLUSTER, as in MODAL, the number of regions can be from 1 to 4. We will alternately refer to these as layers or clusters. In MODAL analyses they have traditionally been called modes as well.

The object of this report is to describe MODAL and CLUSTER, and to compare the results of a large sample of cases where the two algorithms were run side by side. These evaluations are made possible by the AFGL Man-Computer Interactive Data Access System (McIDAS) which is an interactive computer system. The algorithms were programmed for the McIDAS, and a procedure was arrived at for comparing the two algorithms for different cloud situations. After some experimentation, a sample set of 350 cases was selected from smoothed (3 n mi) DMSP imagery. Each case consists of an array of 15 x 16 infrared grayshades, which means that each case contains 4 MODAL analyses, or a total of 1400 MODAL analyses for the entire study. The cases cover a wide range of cloud types, ground types, and combinations of cloud and ground. Considerable insight into the algorithms was gained by image analysts in reviewing these sets of data, and it is felt that the basic samples provide a well-rounded data set not only for this study but for further investigations.

### 2. DETAILS OF THE ALGORITHMS

The basic natures of the two algorithms studied here are very different since MODAL is a histogram evaluation algorithm and CLUSTER is a dynamic algorithm. Details of their structure and operation will now be given.

<sup>3.</sup> Hawkins, Rupert S. (1980) A Clustering Technique for Satellite Image Analysis, Proc. 8th Conf. on Weather Forecasting and Analysis, Amer. Meteor. Soc., 115-118.

<sup>4.</sup> Hawkins, Rupert S. (1981) Objective Analysis of Satellite Cloud Imagery,
Proc. 1981 International Geoscience and Remote Sensing Symposium,
IEEE Geoscience and Remote Sensing Society, Volume 1, 477-482.

#### 2.1 Introduction and Details of MODAL

The IR processor of the 3DNEPH merges infrared data, the surface and upper air temperature fields, and the geography and terrain fields to produce an eighthmesh (25 n mi) cloud analysis on the 3DNEPH grid. Cloud-top heights, temperatures, and total cloud amounts for up to two "most significant" layers are determined. ("Most significant" is defined in the 3DNEPH satellite processor as the two coldest layers separated by 1500 feet or more.)

The automated IR imagery analysis algorithm MODAL is the routine used by the 3DNEPH. MODAL is capable of detecting up to four layers, or <u>modes</u>, of clouds (including "clear" if it is present). It constructs a histogram of the IR grayshades (see Figure 1) of an 8 x 8 IR image array. These grayshades are

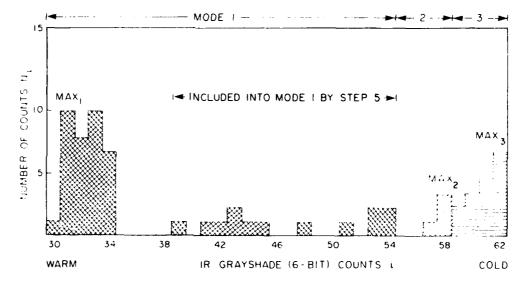


Figure 1. MODAL Histogram Analysis

then separated into modes (there are 3 modes in Figure 1) and a representative grayshade is determined. Using "COLD" as the coldest grayshade within a mode, "MEAN" as the average grayshade of a mode, and "MAN" as the most frequent grayshade within a mode, Table 1 lists the seven possible ways that a mode's representative grayshade is determined for the cloud/no-cloud decision, and for its height. Each mode's representative "height" grayshade value is first converted to an equivalent IR temperature, then corrected for water vapor

Table 1. Options for Representative Grayshades of a Mode

Scheme <u>Number</u>	Cloud/ No Cloud	Height	
1.	COLD	COLD	
2.	MAX	MAX	
3.	MEAN	MEAN	
4.	MAX	COLD	
5.	MEAN	COLD	
6.	MEAN	MAX	
7.	MAX	MEAN	

attenuation and look angle, and finally converted to the height of the mode using the eighth-mesh surface and upper air temperature profile. If the representative "cloud/no-cloud" temperature of a mode is not significantly different from the surface temperature, then that mode is flagged as "no cloud." The total cloud amount is determined by summing the number of pixels in the cloud mode(s) and dividing by the total number of pixels in the image array.

The choice of schemes in Table 1 allows for tuning of the 3DNEPH output to the needs of its users. Currently, the nephanalysis uses Scheme 1. Such a choice tends to overestimate the presence of clouds, since the cold point of a mode may be significantly colder than the MEAN or the MAN of the mode. This scheme also tends to overestimate the altitude of low clouds; however, this scheme may be best for semitransparent cirrus clouds since all of their grayshades may be warmer than the true cloud temperature. An example is given later. A description of the logic MODAL uses to distinguish cloud layers follows

MOD M. operates on the 6-bit IR image values of an 8 x 8-pixel array (25 n mi x 25 n mi nominal size). Its objective is to separate clear regions and cloud layers present in the image sample. The first function of the algorithm is to generate a histogram of the 8 x 8 array (see Figure 1). Modes are then defined in the following manner. MI searching is done from the cold end of the histogram to the warm end, unless explicitly stated otherwise.

<sup>1.</sup> A search is made through the histogram for that grayshade that appears most frequently in the array. This grayshade is the mode of the mode, henceforth called  $\underline{MXX}$  (see Figure 1). The number of grayshades of value MAX must

- be  $\geq 3$ . Let  $N_i$  = # of grayshades of value i. Thus, if  $N_{\mbox{MAX}} < 3$ , go to Step 5.
- 2. Grayshades that are one or two counts warmer or colder from MAX are then unconditionally assigned to the mode that MAX has begun to define; however, if any  $N_{\rm MAX\pm 1}$  or  $N_{\rm MAX\pm 2}$  = 0 (a breakpoint), all further assignment of pixels to the mode on that side (warmer or colder) of MAX is terminated.
- 3. Further inclusion of any more grayshades i into this mode farther away than 2 grayshade values from MAX is strictly dependent on two conditions: (1) the number of pixels Ni must be greater than 0, and (2) the Ni must monotonically decrease with distance from MAX. As soon as either of these two conditions is not met, assignment of pixels to the mode of MAX is terminated. This is done on the warm and cold sides of MAX until either of the above conditions is broken or the end of the grayshade scale is reached, whichever comes first.
- 4. Go to Step 1, ignoring any pixels previously accounted for. Up to 4 modes are allowed. If 4 have been chosen, go to Step 5.
- 5. Go back and include any "leftover" pixels into existing modes if possible. Leftover pixels are those not accounted for in Steps 1-4. Leftovers are placed into the warmest mode that is closest to it (see Figure 1), and in no circumstance will leftover pixels warmer than the pixels in any mode be included in a mode. This ensures that clear pixels will not be inserted into cloudy modes; however, this also allows for some "unassigned" pixels.

There are some undesirable loopholes in this logic that are not immediately obvious. One situation exists when several pockets of IR grayshades with "peaks"  $(N_i)_{i \le 3}$  lie included from a mode so as not to be included in or assigned to that mode. In many cases, single breakpoints separate substantially large groups of pixels from what should perhaps be their own mode (see Figure 1). The 3DNEPH accounts for these pixels in total cloud calculations; however, subsequent layer calculations using the leftover pixels can unjustly increase cloud amounts and heights. Such is the case in Figure 1, Mode 1. The 24-grayshade spread of this mode, representing a 1T of approximately  $40^\circ$  E, is much too great. Although the majority of the pixels in Mode 1 are concentrated around grayshade value 32, the 3DNEPH would choose as a representative value (using Scheme 1 of Table 1) the coldest grayshade in Mode 1, which is 54. The use of such a cold value for

this cloud layer seemingly must overestimate its height. It can potentially overestimate cloud presence as well for modes that have a mixture of clear and cloudy pixels.

Unassigned pixels appeared in 26.7 percent of the 1400 MODAL cases we studied. Most of the time the  $\Delta T$  spread was considerably less than the  $40^{\circ} \mathrm{K}$  in Figure 1 and many unassigned pixels were for cirriform clouds, for which a sizable  $\Delta T$  is expected due to differing degrees of transparency of the cirrus clouds.

Another situation that commonly arises is that of overanalysis. On account of the monotonically decreasing criterion in Step 3, and the breakpoint criterion in Steps 2 and 3, MODAL tends to be too sensitive to small features in the image data. Figure 1 also demonstrates this point well. More logically, Modes 2 and 3 should be considered one mode. In general, MODAL identifies too many cloud modes and misses some clear areas. Some of this may be due, however, to the fact that satellite IR data is relatively more sensitive to high clouds than to low clouds. We have mentioned briefly a few of the apparent shortcomings of the MODAL analyses. We realize nonetheless that AFGWC is aware of these shortcomings and we anticipate that an updated IR image analysis algorithm in the RTNEPH will attempt to solve them. (The Real-Time NEPH\_RTNEPH\_ is a substantial revision of the 3DNEPH scheduled for use in 1983.)

#### 2.2 The CLUSTER Algorithm

CLUSTER was designed to analyze a 16 x 16 army of infrared imagery. Thus four 8 x 8 MODAL-type arrays are contained in one CLUSTER array and, in all the algorithm testing we have done, four MODAL arrays have been calculated for each CLUSTER array. The fact that CLUSTER runs on a 16 x 16 array does not restrict its potential application to the 3DNEPH since the critical grayshades determined by CLUSTER can be easily applied to the four 8 x 8 arrays within the  $16 \times 16$  array to give separate estimates of cloud amounts and altitudes on the eighth-mesh or 25 n mi grid scale.

The first step of this algorithm is to separate the array into four quadrants. Histograms are obtained for sixteen intervals per quadrant. The action of the algorithm is to transform these histograms into clusters. At each step, transfers are made from quadrants having smaller frequencies in an interval to the quadrant having the largest frequency in that interval. This separates the data into unique clusters completely consistent with the overall frequency distribution of the whole area. The critical grayshades, which are the transition grayshades between clusters, are called cut levels.

The algorithm is essentially sequential, and can be stated, therefore, as a list of operations. Small loops sometimes occur within these steps, but no loop

occurs among or between steps. This characteristic makes for very fast execution as compared to algorithms that require many iterated calculations within large loops. The major steps in the clustering procedure follow:

STEP 1. Subscript the 16 x 16 infrared array by quadrants. This dimensioned variable could be labeled IR (4,64). The four is for the four quadrants and the sixty-four for the 8  $\pm$  8 values within each quadrant.

STEP 2. The 3DNEPH grayshades (6-bit or 64 grayshades linear with temperature from 310 to 210 K) are placed in 16 intervals. The widths of the intervals are shown in Table 2, and are recognized to include all or nearly all meteorologically significant data. Interval 1 includes the warmest counts, and Interval 16 includes the coldest counts.

STEP 4. For each interval, place the sum of all four quadrant frequencies into the quadrant having the largest frequency value for that interval. If equals exist for largest, leave for the next step. (Note that the quadrant data are now being transformed to cluster data.)

STEP 4. For those instances with two or more largest frequencies, place the sum at that interval into the quadrant having the largest sum in adjacent intervals.

STEP 5. If any quadrant (cluster) has less than therefore points (5 percent of total coverage) place it in the nearest cluster.

STEP 6. For any cluster consisting of two parts (one or more zero intervals between non-zero frequencies), separate by putting one in a zero cluster if a zero cluster exists. If not, put the one closest to another cluster into that cluster.

STEP 7. Combine close clusters. This closeness value is one of the parameters open in the analysis. It was selected so that clusters consisting of values at only one interval and having a similar cluster one interval away in another cluster are combined into one cluster.

STEP 8. Eliminate clusters with less than 13 points (5 percent of coverage). This is used a second time since the separation of multiple clusters could create a small cluster.

STEP 9. Summarize cluster information; this includes the generation of histograms of clusters, number of clusters, and critical grayshades or cut levels that separate cloud regions from one another.

Table 2. Corresponding Interval Ranges for the 6-Bit DMSP Data Used in This Study. The temperature ranges are determined by linearly converting the ends of the corresponding grayshade intervals to \*K

	DMSP 6-Bit	
Interval	IR Grayshade Range	Temperature Range <sup>o</sup> K
16	52-63	227.3-210.0
15	49-51	232.2-229.0
14	47-48	235.4-233.8
13	44-46	240.1-237.0
12	42-43	243.3-241.7
11	39-41	248.1-244.9
10	37-38	251.2-249.7
9	34-36	256.0-252.8
8	32-33	259.2-257.6
7	29-31	264.0-260.8
6	27-28	267.1-265.6
5	24-26	271.9-268.7
4	22-23	275.1-273.5
3	19-21	279.8-276.7
2	17-18	283.0-281.4
1	0-16	310.0-284.6

These cut levels are the major result of the clustering routine. Along with the infrared values, they specify clustered regions that become the subject of interest, and represent a dimensional simplification of the imagery data. Figure 2 shows some results of calculations for a GOES IR image sample. Quadrant histograms appear on the left, and cluster histograms on the right. Summaries are given below these as marked.

## 3. SUMMARIES OF THE DATA SET

The data set used for this study was taken from nominal 3 n mi resolution Defense Meteorological Satellite Program Operational Linescan System (DMSP OLS) visible and IR data. Only IR data were actually processed, but corresponding visible images were used by the image analysts to help in making accurate subjective cloud amount and type classifications. The OLS is the primary sensor on DMSP spacecraft. It is a dual-channel scanning radiometer that senses reflected light and emitted infrared energy in the 0.4 to 1.0 um and 10.2 to 12.8 um spectral bands, respectively. A more advanced design of the OLS enables the earth sampling to vary much less in resolution as scanning proceeds from nadir to higher scan angles. The 3 n mi OLS data described in this report were smoothed twice: first,

GRAYSHADE	INTERVAL	UNCLUSTERED	CLUSTERED
RANGE	#	QUAD. FREQS.	CLSTR. FREQS.
206 - 255	16	1 1 1 1	1 1 1 1
196 - 205	15	]	
186 - 195	14	3    17	
176 - 185	13	8     38	46
166 - 175	12		1 1 9
156 - 165	11	1 12 1 1 1 3 1	1 1 1 6
146 - 155	10	1 1 3 1 1	4
136 - 145	9	3   5   2	1 10 1
126 - 135	8	1   4	1 1 5 1
116 - 125	7	5 6 12 1 1	1 124 1
106 - 115	6	16   9   9   1	24
96 - 105	5	26   4   18	48
86 - 95	4	122   6   8	361
76 - 85	3	5 117 1 1 1	23
66 - 75	2	1 1 1 1 1	
0 - 65	1	1 1 1 1	1 1 1
	Σ's	64 64 64 64 1	108 24 43 81

CLSTR. NO.	1	2	3	1 4 1
CUT LEVELS	85	_   1	15   1	55
NO. OF PTS.	24	108	43	81
MEAN IR	79.6	98.9	128.6	179.91
MEAN VIS	40.21	53.0	64.0	71.7

Figure 2. Histogram Data Before and Mter Clustering With a Summary Below of the Four Clusters. This example is from an 8-bit IR GOES image

on the satellite from 0.3 n mi to 1.5 n mi, and then at AFGWC from 1.5 n mi to 3 n mi. Take the visible data, OLS IR data are digitized into one of 128 possible grayshade values on the satellite, and truncated to one of 64 possible values at GWC. The IR sensors are designed to detect a maximum equivalent blackbody temperature of 310°K and a minimum of 310°K. Warmer scenes, however, tend to have considerable atmospheric attenuation so that equivalent blackbody temperatures are less than true scene temperatures. Attenuation correction factors vary free 45 at 310°K down to 1° at 210°L.

The data set consists of 350 CLUSTER samples and their corresponding 1400 MODAL samples taken from a wide variety of atmospheric conditions ranging from clear to highly active cloud region. While an effort was made to get as varied a sample as possible, the more prevident cloud types occur more frequently within the set. Also, samples of had-water boundaries under clear skies were selected. The unterlying consideration in a dain, the data selection was to obtain as "complete" a sample as possible for testing algorithms such as MODAL and CLUSTER.

The samples are mostly from regions of feature transition. Flat regions were occasionally selected but, since these do not generally reveal much about the capabilities of the algorithms, samples that include more than one level or feature (that is, land/cloud) were ordinarily chosen. These levels range in contrast from multilayered cloud scenes to simpler clear/cloud scenes.

The data set was sent on tape by AFGWC to AFGL, and eighth-mesh samples were collected and saved using the McIDAS. Data and algorithm results were printed out for each of the 350 cases in the set. The results for Case 302 are shown in Figure 3. Data such as shown in Figure 3 was used to compare the performance of the algorithms. At the top of Figure 3 some information concerning which imagery the case came from is presented. The data, cloud types observed, and background (which were entered by the operator during the sample save generation) are given. The top array contains the original 6-bit infrared image values. (A value of 0 is warmest, and 63 is coldest.) Results of MCDAL and CLUSTER are then listed in coded form. "L" represents the lowest layer (warmest mode or cluster),"." the next highest (next coldest) layer, "-" the next, and "H" the highest layer. (Note that if only two layers are found, "." represents the highest [coldest, one.) Grayshade ranges for the four MODAL arrays and one CLUSTER array are then given at the bottom of the printout.

Figure 3 also illustrates how the CLLSTER algorithm would easily generate a cloud analysis for 8 x 8 IR arrays (eighth mesh) or even 4 x 4 arrays (sixteenthmesh). The critical grayshades or cut values can be determined for the 16 x 16 array, as CLASTER already does. These cut values can then be applied to compute cloud parameters such as cloud cover estimates on  $6 \times 6$  or  $4 \times 4$  pixel subsets of the original  $16 \times 16$  arr c.

The example in Figure 3 is for Sample 302 which, according to the operator's notes recorded at the time of the save, consists of about 50 percent cumulus buildings and 50 percent stratus buildings. Remember that MODAL's smallyses are for four independently analyzed 8 x 3 arrays, all of which make up the 50 percent Cu and 50 percent 8t, and are merely positioned together in Figure 3 for comparison with the CLUSTIR analysis. This causes some eventual irregularities in the layers' representative IR may shades to occur from one 3 x 3 bex to another, whereas a smooth analysis is sent inted by CLUSTIR. Any such small-scale irregularities will be reflected in the cloud analysis output.

## 3.1 Categories of Clear and Cloudy Regions

The 350 cases were separated into seven basic categories. Table 3 shows the names of these categories along with the number of cases. From left to right

```
9447_F #312
                                     MOJAL-CLUSTER ANALYSIS 6-BIT IR GRAYSHADES
        1443E 38102+130
08811 #2751
                                 OFC. 17, 1979
GLOUDS: CA ST
        TV _TN 512
TV ELF 364
                                 SKGND: 4AT
                27 29 20 29 30 31 32 34 34 35 35 39 39 36 31 30 30 20 20 20 32 31 32 34 34 35 39 41 40 35 30 30 30 30 20 24 32 33 31 32 33 35 38 40 42 39 34 50 30 30 30
                32 24 54 55

52 24 54 55

54 51 54 55

55 57 54 55

56 57 54 55

58 40 54 55

58 40 54 55
                                30
29
30
                   3 n
3 0
                   #77AL AVALYSIS
            + " ; AL AVALYSIS
                                                                  CLUSTER AVALYSIS
                                    MODE 2=.
RNG CNTS
30+52 14
34-36 12
34-41 31
33-34 7
                                                                 MODE WEH CLUSTER
RNG CNTS # RNG CNTS %
39-40 5 144 0-31 64 25
42-43 4 25, 32-36 41 31
0-2 0 32- 37-41 82 36
                                                   470E 3==
4NG ENTS
33=38 35
37=41 19
25
31
                                                   42-44 15
```

Ti use 3. MODAL and CLUSTER Analysis for Case 302

these are Clear Land/Water/Snow, Cumulus, Stratus, Stratocumulus, Cirrostratus, Cirrus, and Cumulonimbus.

Table 3. Composition of the Cloud Imagery Sample Set

1								
	CLL/CLW/CLS	Cu	St	Se	Cs	Ci	СЪ	
	48	85	45	39	16	92	25	
		. J i		l. :	l	l	i i	i

This separation is done on the basis of the major classification (the most dominant type) for each sample. Minor constituents usually occur in all cases except clear land or clear water. We feel that this sample is more than adequate for evaluating MODAL and CLUSTER-type algorithms. The fundamental drawback in the evaluation of such algorithms is the subjective approach that is required to settle on (a) what is desired of an algorithm like MODAL, and (b) once decided upon, how well MODAL and other algorithms meet those criteria. The criteria we feel to be of importance are enumerated in the following paragraphs. Also stated is the rationale for separating a grid box into regions or layers.

Separation of basic features for making simple accurate statements about grid boxes is the goal. In the 3DNEPH these features (regions) are referred to as layers. Naturally, layers standing out from other layers are most desired to be separated. Two layers hardly distinguishable from each other should be analyzed as one layer. Both the size of the area and how much the area stands out from adjacent areas should determine whether or not it qualifies as an independent layer. These goals are not stated explicitly; however, they are implicitly included in algorithms that perform the separation.

The separation should be mutually exclusive. In other words, the sum of all the layers/clusters should constitute the whole grid box area. Among other things, this makes for a complete analysis and more meaningful data reduction than if this is not a requirement. CLUSTER has this feature. MODAL lacks this feature since sometimes IR values are not assigned to modes; however, subsequent processing in the 3DNEPH decides whether the unassigned IR values are clear or cloudy.

The sensitivity of the separating algorithm should be reasonable from the warmest equatorial regions to the coldest polar regions. It would also be an

advantage if it could be tuned to give a desired response. Both MODAL and CLI STER are open to tuning, but CLUSTER is far more flexible in this respect.

For operation in the 3DNEPH model, fast computer times are required to perform the analysis due to real-time processing constraints. If an algorithm is not fast enough it cannot qualify for use by the 3DNEPH. Of course, MODAL satisfies these requirements. Our experiments with run times show that, on an average, MODAL takes 0.026 seconds, and CLUSTER takes 0.094 seconds. Comparing area-weighted times, CLUSTER runs a little faster than MODAL does.

### 3.2 Evaluation of the Two Algorithms in Terms of the Seven Basic Categories

Clear. From the point of view of region separation, this is a simple category. Problems arise not from algorithm inadequacies, but from the very nature of the IR data. There are instances when low clouds blend into the surface temperatures and, consequently, are not separated by either MODAL or CLUSTER. Therefore, clear regions and scattered cumulus conditions are often difficult to separate satisfactorily. Figure 4 illustrates how the 16 x 16 CLUSTER array size has a "context" advantage over MODAL, when scattered cumulus are present in an otherwise clear area. CLUSTER identifies two distinct layers, compared to just one layer for MODAL.

In warmer regions where land temperatures approach those of water, subtle land/water distinctions are barely discernible by the IR sensor, and commonly indiscernible by the MODAL and CLUSTER algorithms (Figures 5 and 6). Wherever temperature contrasts are great enough, neither MODAL nor CLUSTER has serious problems in land/water separations.

Cumulus. There are several problems with the ability of both MODAL and CLUSTER in the analysis of Cu. Most of these problems, however, are due to the physical nature of the clouds themselves. They generally do not fill the field of view of the sensor. Thus, calculations of total cloud are difficult to perform, especially on a pixel-by-pixel basis as the 3DNEPH essentially operates. An IR pixel that represents a Cu cloud that partially fills that pixel contaminates the grayshade: the temperature that grayshade represents is actually representative of some cloud and some ground radiation. The observed temperature is warmer than the actual cloud element's temperature, but colder than the underlying surface temperature. When the algorithms compare such a temperature to an adjacent clear pixel, the difference in temperature (grayshade) may or may not be enough to allow MODAL or CLUSTER to distinguish between a partially cloud-filled region and a clear region. It should be noted that this is not necessarily an algorithm problem, but a physical one due to IR sensor characteristics, field

```
5443_E +157
                         VIDAL-CLUSTER ANALYSIS 6-BIT IN GRAYSHADES
     TMASE 24002-123
0851, $150
0851, $15
0851, $15
0851, $15
                      DEC. 25, 1979
CLOUDS: CJ
BAGNU: LAN
                                  CLL
          15 15 15 17 17 13 13 15 17 19 16 15 16 15 16
43346 44461515
                                    CLUSTER AVALYSIS
                        7076 25.
445 0418
14 1 0
04 1 1
                table 1=c
                                  4))E 5=-
                                            430E 4=4
                                  HNS CNTS
U- U 0
U- 3 0
         17:5 443 C+15
                                           443 2478
0- 0 0
                                                 0 1=L 0-10131 51
0 2=. 19-63125 48
1 1 TVANCALL
               17-21 64
                                            0 = 0
0 = 0
                               )
                                                  0 3=-
S INAPORUL
                                                        0 - 0 u
           1
1 INAPCALE
               12-15 52
```

Figure 4. MODAL and CLUSTER analysis for Case 167

```
MODAL-CLUSTER ANALYSIS 6-817 IR GRAYSHADES
    SAWPLE #349
    1443E 30002-120
                  DEC. 16, 1979
CLOUDS: CLL CLM
RKGND: MIX
    04811 #2737
    TV _IN 373
          12
        12
        12
          12
        12
                      11 11 11 11 11 11 10 10 11 11
         11 11 11 11 11 10
                  11 11 11 11 11 10
                                10 11
           11 11 11
                                       11 11 11
               10
                  11 11
                       11 11 11 10
                                11 10
                                     11 11 11
           11 10 10 10 10 10 11 11 11 10 10
                                     11 11 11
          11
                                            11
          *JOAL AVALYSIS
                                    CLUSTER AVALYSIS
# MODE 1=L
MODES RNG CNTS
1 12+13 64
                    400E 2=.
                            MODE 3=+
                                              CLUSTER
RNG CNTS
                                    7N3 CNTS
                    0-0 0
0-1 0
                            RNS CVTS
                                            .
L TYAPCALL
                                         0 1=1
                                               1-63256 100
                                     0- 0
                                         0 3=-
                                               1= C
0= 0
THARCALT
       5
         1
            11-12 64
                             0 +
                               า
                                                    J
                                                       ٥
SUADRANT 3
         1
            10-12 64
                     0 - 0
                          ĵ
                             n = 0
                                  0
                                                       0
                             0 -
U TZAHCALE
            10-11 54
                     0 - 0
                               0
```

Figure 5. MODAL and CLUSTER Analysis for Case 340

```
MODAL-CLUSTER ANALYSIS
                                                                 6-91T IR GRAYSHADES
      SAMPLE # 50
      IMAGE 35002-130
ORRIT #2690
TV _UN 361
TV ELE 185
                             DEC. 12, 1979
CLOUDS: CLL CLW
BKGND: MIX
             MODAL AVALYSIS
                                                           CLUSTER ANALYSIS
    -----
       .....
1111
                                                                                        L
                                                                                        1
    .....
       . . . . . . .
                                                                                        L
    111
       L
                                                                                        L
  L
            MODE 1 #L
MODES RNG CNTS
2 12-14 58
1 12-15 64
1 13-15 64
                                 MODE 2=.
                                              ♥0)E 3=-
                                                           H=B 30CF
                                                                            CLUSTER
                                 MODE 2=.

ANG ENTS

15-15 6

n= 0 0

0- 0 0
                                                          NODE WELL
RNG CNTS #
0+ 0 0 1=L
0+ 0 0 2=.
                                              9N3 CNTS
0- 0 0
                                                                             RNG CNTS X
                                                                             0-63255 100
1 THANCAUS
                                                            0 - 0
                                                                             0 - 0
                                                                                    0
JUADRAVI 5
                                               ე -
                                                   0
                                                                    0
                                                                      3=-
                                                                             0 - 0
                                                                                    υ
                                                                                         ٥
                                      ũ
```

Figure 6. MODAL and CLUSTER Analysis for Case 60

of view, and atmospheric attenuation. Nonetheless, CLUSTER seems to handle these situations more favorably than MODAL, since MODAL "tries too hard" to separate subtle changes in grayshade. Separation of Cu from higher layers does not present a problem for either algorithm.

Figures 7 and 8 illustrate the action of the two algorithms on cumulus fields. First, notice that MODAL overanalyzes (provides seemingly irrelevant detail) in some places, and that CLUSTER gives an apparently neater analysis. Points not analyzed into a level by MODAL are blank in the illustrations. These points are warmer than the warmest mode (see Step 5 of MODAL in Section 2.1). While the two cases represent specific situations that are not that frequent in the data samples, they do indicate MODAL's apparent problem of overanalyzing the image data. MODAL might be characterized as being "oversensitive" to small changes over small distances. Some of the variability in the MODAL analysis occurs at the boundaries of the quadrants. MODAL is not at fault here since the individual quadrants may have appropriate layers. It must be remembered that the MODAL mapping is made up of four independent runs. Symbols may not represent the same temperature range from one quadrant to the next.

Stratus. Neither MODAL nor CLUSTER has a problem with flat fields of data. Stratus is one cloud type on which the two algorithms behave similarly. There are cases to be found throughout the stratus samples that support this observation; however, again there is a tendency for MODAL to be oversensitive, and put in more layers than necessary.

Stratocumulus. Stratocumulus poses less of a problem than cumulus since these clouds tend to fill more of the field of view. Thus, many of the physical constraints inherent in cumulus detection and analysis are not so prevalent for the proper detection and analysis of Sc. Both MODAL and CLUSTER tend to find too many layers of Sc. They do not always do so on the same cases. Examples of MODAL overanalysis of Sc are in Figures 9, 10, and 11, and for CLUSTER in Figures 12, 13, and 14. Both are finding subtle changes in the temperatures of cloud fields. We feel that the cloud analysis would be improved if this sensitivity was tuned out of the algorithm to be used in the 3DNEPH.

<u>Cirrostratus</u>. Cirrostratus fields tend to appear flat and are generally well-analyzed, since temperature differences among these ice clouds and most any other categories in an image are great. While both algorithms perform well, MODAL has the tendency to find too many layers, as noted in other categories (see Figure 15, for example).

```
S442_E #312
                                      MODAL-CLUSTER AVALYSIS 6-BIT IR GRAYSHADES
        14435 20002-110
        URBIT #2737
TV _1V 31+
TV ELE 350
                                 DEC. 16, 1979
CLGUDS1 CJ CLL
                                  BAGND: LAY
                16 15 14 14 14 15 17 27 35 40 43 48 49 41 40 24 14 15 15 14 15 15 27 35 43 43 46 50 53 52 36 13 15 16 14 16 16 16 24 35 45 47 48 52 54 54 52 15 13 14 14 15 17 17 21 32 46 50 50 50 53 53 49 15 15 13 14 15 16 17 21 30 40 40 52 55 56 55 52
                15 15 15 15 15 15 15 17 17 25 30 49 53 56 56 56 15 15 15 15 16 17 23 51 41 40 54 55 54 14 14 14 14 15 15 15 15 15 16 17 23 30 40 43 43 47
                                                                                 53
52
45
                 14 15 14 15 13 14 14 15 16 17 24 27 56 42 46
                 13 13 13 14 14 13 13 14 15 16 15 25 26 29 29
                 13 13 13 15 15 13 14 14 14 15 16 20 23 29 34
                13 13 13 13 13 13 14 14 14 14 15 16 22 26 27 24 15 16 15 15 15 17 27 21 21 15 14 14 14 15 15 16 12 26 27 24 15 16 15 15 17 27 21 21 15 14 14 14 14 14 14 15 15 15 16 16 19 22 26 13 14 14 14 14 15 15 16 16 19 21 23 23 21
                05 55 55 15 15 15 16 16 16 22 22 21 22 22 20
                                                        4 104 4 44 4 415
                                                         400E 1=0
                                                     40)£ 3=-
                                                                   RNG CNTS # RNG CNTS 4
               #1765 445 0415
1 15-27 44
                                                    HY CYTS
                                      HNG CN19
JUANKAVI 1 1 JUANKAVI 2 u
                                     45-30 15
                       45-45 5
                                                                   55-56
                                                                           5 2=. 19=26 35
0 3== 27=63 73
                                                    52-35 15
0- 0 n
                1
                                                                    0- 0
                                                                                                    28
                                                    26-27
1 IF APCALE
                                     14-25 23
```

Figure 7. MODAL and CLUSTER Analysis for Case 312

```
MUDAL-CLUSTER ANALYSIS 6-811 IR GRAYSHADES
       24+3 E #215
       1443E 30002-120
       URBIT #2737
1v 11v 257
Tv ELE 215
                            DEC. 16, 1979
CLOUDS: CJ 51
                            HAGNO: NAT
             43742 444ETSIS
                                                        CLUSTER AVALYSIS
                                                       HIJDE 4EM
HIG CNTS # RNG C
0-0 0 1EL 0-381
0-0 0 2E. 59-65
59-59 4 5E- 0-0
                    4375 1=L
                                            400E 3=-
                                MCDE 2=.
                                                                       CLUSTER
                   10-25 9
23-26 7
31-34 10
                               443 2418
24-25 15
27-37 49
 5 1 1 4 APCAUL
5 1 1 4 APCAUL
5 5 TARPORUS
                                           24-54 42
58-42 8
                                                                        RNG CNTS
                                                                        0-38163
                               27=51 4
35=43 25
-38 4
             - 5
TVANCAUL
TVANCAUL
TVANCAUL
                                           50-56 20
                                                                               0
                                                                                    0
              •
 T. AHCALL
                                45-45
```

Figure 8. MOD M. and CLISTER Analysis for Case 342

· And the second second

The state of the s

```
SAMPLE # 15
                                             MODAL+CLUSTER ANALYSIS
                                                                                       6-BIT IR GRAYSHADES
         14A3E 36032-130
0RBIT #2623
1V LUN 131
VV ELE 283
                                        DEC. 8, 1979
CLOUDS: SC
BKGND: WAT
                                 23 21 22 23 23 21 24
24 23 21 24 24 21 28
21 22 21 27 29 22 24
21 23 24 26 22 21 23
                    25 27 28 28
30 30 30 31
31 31 31 31
                                                                                           21 23
24 26
30 28
30 30
29 28
25 26
27 29
29 30
27 24
27 28
25 27
                                                                                     21
26
29
                             31 30 28
30 26 27
27 29 29
27 27 29
                    31
                        31
30
25
29
30
                   31
28
28
29
30
                             30
                                                                                      56
                            30 30
31 29
27 28
31 26
27 27
29 26
29 24
                        30
30
33
                                                                                      30
                    31
30
30
                                                                                      28
CLUSTER ANALYSIS
               4374L AVALYSIS
                  # MODE 1=L
MODES RNG CNIS
                                                              400E 3=-
                                                                               #3)F 4=H
81V3 2VF
0 0 =0
                                             40)E 2=.
                                                                                                     CLUSTER
                                            PNG CNTS
                                                              RNG CHTS
                                                                                                      RNG CNTS %
                           25-31 60
20-23 37
20-23 16
19-21 9
                                                                                          0 1=L 0-23 70
4 2=. 24-28 99
0 3=- 29-63 87
1 THARCAULS THARCAUL
                  1 2 3
                                                                                                                      27
38
                                                      a
                                                              25-27 13
                                                                               30-30
                                            24-24 10
                                                              25-30 42
```

Figure 9. MODAL and CLUSTER Analysis for Case 15

```
MODAL-CLUSTER ANALYSIS 6-BIT IR GRAYSHADES
     SAMPLE # 16
     IMAGE 36002-130
ORBIT #2623
                       DEC. 8, 1979
CLOUDS: SC
     TV _IIN 132
                       BKGND: MAT
                                               29 25 19 27 27 20 18 22 17 16 16 15 17 17 17 19
           MODAL ANALYSIS
                                             CLISTER AVALYSIS
                         22+22 3
21-25 18
          MODE 5=-
                                              H=# 300F
                                                          CLUSTER
                                              RNS CVT8 # RNG CNTS % 0- 0 0 1= 0-19 88 34 0- 0 0 2= 19-23 56 21
                                   RN3 CNTS
                                             RNS CHTS
                15-20 33
S TVANCALL
                                   24-30 28
                14-20 34
                                   26-29 12
E TVANCALL
                                              0 - 0
                                                      3=- 24-63112
TVASCALL
                          20-23 19
                                             25-28
                                                           0 = 0
```

Figure 10. MODAL and CLUSTER Analysis for Case 16

```
MUDAL-CLUSTER ANALYSIS 6-317 IR GHAYSHADES
                                   SAMPLE #217
                                  1443E 40012-131
0841T #2755
                                                                                                                                               DEC. 18. 1979
                                   TV _IN 341
TV ELF 441
                                                                                                                                                CLOUPS: SC CLA
                                                                                                                                                BEGNO: MAT
                                                                       29 29 27 24 23 24 23 24 25 18 23 27 28 29 25 17 30 27 22 23 27 29 29 29 27 26 23 20 21 26 29 26 30 31 31 21 20 26 27 23 22 28 29 25 21 23 25 29 28
                                                                                                                                                                               31 21 20 26 27 23
29 24 19 21 21 18
29 29 27 21 16 19
20 29 27 19 21 19
19 22 21 19 21 23
16 19 22 26 25 20
25 19 26 29 26 20
24 25 27 28 26 19
28 24 27 23 19 17
28 24 27 23 28 24
                                                                       20
                                                                        18
                                                                       26
                                                                                     28
SURTER AVALYSIS
                                                      4004L AVALYSIS
                                                              # 400F 1=L

400F 8 HNG CNTS

1 2 16=25 42

2 17=21 A

10=17 4

19=22 9
                                                                                                                                                              000 2=.

000 2005

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

000 200

00
                                                                                                                                                                                                                                                                                         #00F 4=H CLUSTER
RNG CNTS # RNG CNTS t
n= 0 0 1=L 0=23 75 29
J= 0 0 2=. 24-26 52 20
25-30 42 3== 27-63129 50
                                                                                                                                                                                                                             403E 3=-
403 CVTS
0= 0 0
   T S INVECTOR
                                                                                                                                                                                                                                                                                                                                                                                                                                     29
                                                                                                                                                                                                                             26=29 38
25=25 6
0= 0 1
    JUASPANT 4
    AJADRANT .
```

Figure 11. MODAL and CLUSTER Analysis for Case 217

```
S442_F # 52
                                MODAL-CLUSTER ANALYSIS 6-BIT IR GRAYSHADES
       1443E 35002=130
DWHIT #2640
TV LIN 275
TV FLE 302
                             DEC. 12, 1979
CLOUDS: SC
                             SKGND: MAT
             477au avairsis
                                               CLUSTER AVALYSIS
           ADDAL ANALYSIS
                               MODE 2=.
RNG CVTS
n= 0 n
n= 0 n
n= 0 n
            # MODE 1=L
                                            400E 3=-
                                                        43DE 4=H
                                                                        CLUSTER
                                            RN3 CNTS
0- 0 0
0- 0 0
0- 0 0
                                                        2N3 CNTS
2-0 0
                                                                         RNG CNTS
                                                                    ø
1 TVANCAUL
5 TVANCAUL
6 TVANCAUL
1 TVANCAUL
                                                                                   %
47
                   50-54 84
50-52 84
50-59 84
              t
                                                                   1=L
                                                                         0-21122
                                                         0- 0
                                                                  2=. 22-63134
3=- 0= 0 0
4=H 0= 0 0
              1
                                                                0
                                                                                     0
```

Figure 12. MODAL and CLUSTER Analysis for Case 62

```
MODAL-CLUSTER ANALYSIS 6-BIT IR GRAYSHADES
     SAMPLE #209
      IMAGE 40002-130
                        DEC. 18. 1979
CLOUDS: SC | BKGND: SIX
      0881T #2755
     TV _IN 57
TV ELE 337
                                    CLL CL*
           CLUSTER AVALYSIS
                407E 1=L
                           400E 2=.
                                     400E 3=-
                                               MODE 4=H
                                                             CLUSTER
                                                     75 # RNG CVIS
5 1=L 0-28 79
7 2=. 29-33 78
0 3=- 34-63 99
7 4=H 0+ 0 1
                                     RNS CHTS
                                               PVS CVTS
                24-29 42
                           $0-30
                                7
                                     32-34 10
                                               35+36
            ن
خ
ک
                                                0 = 0
0 = 0
S TRANCAUL
                 28-29 5
                           30-33 17
                                     34-38 42
                 28-31 21
26-29 35
                           32-36 43
30-36 29
                                      0- 0
                                           0
T. INARCALE
```

Figure 13. MODAL and CLUSTER Analysis for Case 209

```
SA49.E #301
                                        MODAL-CLUSTER ANALYSIS
                                                                             6-BIT IR GRAYSHADES
        1443E 38002-130
        TV _IN 320
TV _IN 320
                                   CLOUDS: SC
BKGND: MAT
                 32 33
31 33
31 32
32 32
CLUSTER AVALYSIS
             MUDAL AVALYSIS
                                                                     4) TE 424 CLUSTER

913 CSTS # AND CSTS $

0- 0 0 1=L 1-28 58 22

0- 0 0 2=. 24-31 95 37

0- 0 0 3=- 32-63102 39
               # MUDE 1#L
27-34 64
27-34 64
27-34 64
27-35 6
37-29-40
                                                      MUDE 3=-
RN3 CNTS
0 - 0 0
                                       400E 2=.
                                       4NG CNTS
0- 0 0
S S JARRCAUC
S S JARRCAUC
S S JARRCAUC
S IVARCAUC
S IVARCAUC
                                       28-35 55
30-32 24
29-53 50
                                                                000
                                                        0- 0
```

Figure 14. MODAL and CLUSTER Analysis for Case 301

```
MODAL-CLUSTER ANALYSIS 6-911 IR GRAYSHADES
         SAM3_E #144
         IMAGE 31002-120
                                    DEC. 18, 1979
CLOUDS: CS CJ
BRGND: AAT
         JRH11 #2755
         Tv _IN 271
Tv =_E 422
                   14 44 52 52 51 50 52 52 51 47 44 39 42 46 47 49
                  45 45 52 52 49 48 49 50 51 49 47 44 43 45 48 48

45 45 45 49 47 46 47 45 51 49 48 49 46 45 45

46 47 44 47 48 48 47 47 49 49 49 50 50 47 45 47

42 47 42 46 48 50 47 47 47 48 48 49 50 46 48 45
                  47 47 45 41 45 47 47 47 46 47 48 49
                   45 44 44 45 45 45 45 41 44 45 45 48
                  19 45 51 46 46 47 45 45 41 39 39 44 36 41 30 25 50 45 50 46 40 37 42 42 45 42 51 27 26 25 23 22 23 35 45 50 44 36 36 42 42 45 42 51 27 27 25 23 22 22 23
                      55 35 35 35 36 42 34 25 24 25 24 22 23 25
                  CLUSTER ANALYSIS
CLUSTER AVALYSIS
             MUJAL AVALYSIS
                                                                        4006 48H CLUSTER
RYG CYTS # R45 CYTS %
0+ 0 0 1=L 0-26 68 26
46-51 37 2x. 27-38 41 16
45-50 7 3x- 39-5147 77
0- 0 0 4xH 0= 0 0 0
 # AUDE 1=L ADE 2*.

210 244 277 245 2408

210 244 2 4 45-51 55

21 1445(AU)

2 4 54-54 5 41-42 7
                                                         4036 5#4
445 5479
52452 5
                                                         45-45 15
 S THANCAUL
                          25-30 25 52-40 12
                                                         42-45 14
                                          n = 3 0
                                                          0-10
                          22-12 64
```

Figure 15. MODAL and CLUSTER Analysis for Case 199

Cirrus. Both algorithms tend to find too many layers of cirrus. MODAL's over-analysis appears to be worse than CLUSTER's, however. Cirrus boundaries naturally are the cause for this problem, as transition zones among solid cirrus and other regions are partially filling the sensor field-of-views. Also, cirrus variations in thickness and, hence, radiative temperature contribute to this problem. While some altitude variation of cirrus is expected, the algorithms respond too strongly to field-of-view and emissivity variations. As with previous cases, these sensitivities should be tuned out of the operational algorithm.

Cumulonimbus. Both algorithms have little problem detecting the presence of cold regions adjacent to warm regions; however, both algorithms tend to be too sensitive to Cb edges. MODAL seems to find more layers at the Cb edges (Figure 16) than CLUSTER. It is interesting to note that in many Cb cases where MODAL has "undefined" pixels (see Step 5 of MODAL in Section 2.1), clouds are generally found. In addition, these areas seem to be where CLUSTER "overanalyzes."

#### 3.3 Frequencies of Clusters and Modes

The 350 cases were reexamined to study the tendency of MODAL and CLUSTER to find too many cloud layers. For this study, we looked at clear cases that were entirely clear, and cloudy cases with only one category of cloud. We were left with 47 clear cases: 27 with a mixture of land and water, and 20 with a uniform background. We also found 168 suitable cloudy cases. The cloudy cases could be overcast or partly cloudy, as long as the background was uniform and only one cloud category was present in each case. The objective was to identify cases where the McIDAS observer, using both IR and visible imagery, found only one or two regions per case, so that only one or two clusters or modes would be the desired result of the algorithms.

Figure 17 shows the percent distributions of the numbers of clusters and modes for these cases. Cirrostratus and cumulonimbus are not shown since there were too few cases to define a distribution. CLUSTER performed well for cumulus cases since the desired one or two clusters were found 86 percent of the time. For clear cases, both MODAL and CLUSTER performed well. There were a few cases over cold land or snow where CLUSTER found two or three clusters when only one was desired. The larger array size for CLUSTER may be more sensitive to thermal gradients over clear land than the smaller array for MODAL. We will look at this in future studies to ensure that the colder land clusters are not improperly identified as clouds.

```
SAMP_E #347
                                                                                                                                                                                                 MUDAL-CLUSIER ANALYSIS 6-BIT IR GRAYSHADES
                                        1-43E 50072-120
UNDIT #2737
TV _1 V 257
TV ELE 257
                                                                                                                                                                        DEC. 16, 1979
CLOUDS: C3 CI
BKGND: MAT
                                                                                                                                                                                                                                     52 52 51 50 59 50
52 52 52 52 52 52
52 52 52 52 52 62
52 52 52 52 52 52
52 52 52 52 52 52
51 50 51 52 52 52
51 50 51 52 52 52
51 50 51 52 52 52
                                                                                                                                                                                                                                                                                                                                                                     61 62 62
62 62 62
62 62 62
62 62 62
62 62 62
                                                                                   56 95 95 96 95 95 95
85 95 95 96 95 96 95
85 95 95 96 96 96 95
                                                                                61
                                                                                                                                                                                                                                                                                                                              62 61
53 54
34 41
                                                                                                                                                                                                                                                                                                                         36 30
38 38
25 31
26 28
24 30
24 29
                                                                                                                                                                                                                                                                                                                                                                                                                    46
                                                                                                                                                                                                                                                                                                                                                                        30
37
33
33
30
31
                                                                                                                                                                                                                                                                                                                                                                                                                     31
                                                                                                                                                                                                                                                                                      CLUSTER AVALYSIS

THE HERE AND 
#3)4( Ava(rs18
                                                              43346 44464515
  7375 1= 1

7170 204 27CB

61 42-52 5 1 1/AHCAUK

76 52-75 1 5 1/AHCAUK

77 55-75 1 1/AHCAUK

78 55-75 1 1/AHCAUK

78 55-75 1 1/AHCAUK
                                                                                                                                                                                                                                                                   MODE 3=-
H += CNTS
0= 0 0
                                                                                                                                                                                                                                                                                                                                           MUDE 4=M CLUSTER
443 CNTS # RNG CNTS %
0-0 0 1=L 0-28 39 19
0-0 0 2=. 29-38 59 2
                                                                                                                                                                                          MIDE 25.
RAU CYTS
51+52 41
J+ 0 3
                                                                                                                                                                                             29-23 11
                                                                                                                                                                                                                                                                    30-35 26
                                                                                                                                                                                                                                                                                                                                            54-49 19 3=- 39-46
                                                                                                                                                                                              53-50 10
                                                                                                                                                                                                                                                                    28-34 22
                                                                                                                                                                                                                                                                                                                                           35-62 24 424 47-65130
```

Figure 16. MODAL and CLUSTER Analysis for Case 349

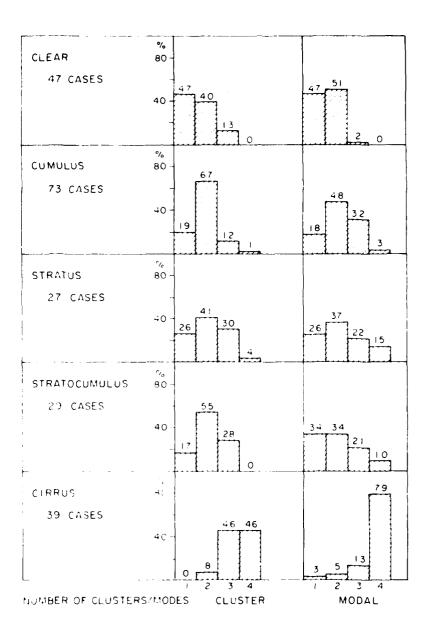


Figure 17. Percent Distribution of the Numbers of Clusters/Modes for Entirely Clear Cases and Cloudy Cases With Only One Category of Cloud Per Case.

The cloudy categories in Figure 17 confirm the examples in the previous section. Both MODAL and CLUSTER are tending to overanalyze cloud layers, finding 3 or 4 modes or clusters when only 1 or 2 are desired. Moreover, MODAL is overanalyzing more frequently than CLUSTER. The cirrus cases are the worst. (The few cases of pure cirrostratus and cumulonimbus also tended to have 3 or 4 modes or clusters.) It can be argued that cirrus may be found in multiple layers so that 3 or 4 clusters or modes would be acceptable for some cases. Nothing in our meteorological experience, however, leads us to believe that this should happen 92 percent of the time as Figure 17 shows when the percentages for 3 and 4 clusters/modes are summed.

Two further comments related to the excessive number of modes found by MODAL should be made. First, the most recent technical report on the 3DNEPH states that a mode is defined for a single grayshade that contains 6 or more samples, or a group of adjacent grayshades containing 12 or more samples. This is no longer the case in MODAL, since a single grayshade that contains 3 or more samples can presently define a mode. These "small" modes contribute heavily to MODAL's tendency to find more modes than observers can find in the IR or visible imagery. The second comment is that not all the modes in MODAL are designated as separate cloud layers by the 3DNEPH. There are somewhat arbitrary rules elsewhere in the 3DNEPH code for combining the modes into no more than two cloudy layers. Although we did not duplicate these rules in this study, we suspect that the overanalysis of modes will still be detrimental to the final cloud layers in the 3DNEPH, at least when cirrus is present, and probable for other cloud categories as well.

As anticipated, both algorithms performed extremely well (over 90 percent accuracy) when analyzing samples where only one mode or cluster was desired. It should be noted, however, that the majority of cases containing only one homogeneous region were clear land, water, or snow and uniform stratus, imagery types that present only minor problems to the algorithms.

## 3.4 Preliminary Cloud Cover Evaluation

While cases were being identified and saved on the McIDAS, the image analysts made subjective estimates of the fraction of cloud categories over the 16 x 16 array of IR values. The observer could look at both visible and IR images so that the estimates of cloud cover for low clouds were generally trustworthy, and were assumed accurate to #10 percent. As in the previous section, "simple" cloud cases with only one cloud category appearing in each case were studied. An observer reviewed each case for both MODAL and CLUSTER, identified the cloudy modes and clusters, and calculated the total cloud cover for each 16 x 16

array based on the numbers of samples in the cloudy modes and clusters. The observer was performing the role of the 3DNEPH surface temperature field in separating the clear and cloudy modes or clusters. (This was a time-consuming technique that we will endeavor to improve for future cloud analysis studies on the McIDAS.) Results are given in Tables 4 and 5.

Table 4. Cloud Cover Within 10 Percent of Observed

· · · · · · · · · · · · · · · · · · ·			
	CLUSTER	MODAL	Total Cases
Cu	38%	25%	52
St	61%	70%	23
Se	80%	60%	25
Ci	50%	38%	34
All Four Categories	532	43%	1 34

Table 5. Cloud Cover Closest to Observed

	CLUSTER	MODAL	Ties	Total Cases
Cu	46%	42%	12%	52
St	30."	13%	57%	23
Sc	440	28%	287	25
t i	385	47.	15%	34
All Four Categories	410	36%	23%	134

Table 4 shows the percentages of cases that had cloud cover within 10 percent of the observed. For the total of the four categories of clouds, CLUSTER performed better than MODAL (53 percent compared to 43 percent) but there is obviously room for improvement for both algorithms. The cloud cover estimates were weakest for cumulus and cirrus categories, which is expected since these categories are most likely to have subpixel cloud elements and highly variable changes in cloud emissivities from one pixel to the next. The categories with

larger cloud elements, St and Sc, had more accurate cloud cover estimates for both algorithms.

Table 5 shows the percentages of cloud cover estimates that were closest to the observed values on a case-by-case basis. CLUSTER had cloud cover estimates closest to observed for 3 out of 4 categories, and for the sum of all categories; however, the differences between MODAL and CLUSTER are not great in this comparison.

The average cloud covers for observed, MODAL, and CLUSTER estimates were calculated for each category to see if the algorithms were finding more or less clouds than the observed. Very little difference was found for Cu, St, and Se, but MODAL and CLUSTER both found about 5 percent more Ci than the image observers. The image observer estimates of Ci coverage often seemed low compared to hund-drawn contour analyses of the 16 x 16 arrays of IR data. For this cloud category, moreover, the visible data was not very helpful for estimates of cloud cover from the image displays.

The subjective estimates of percent cloud cover were almost indispensable for evaluating the performance of CLUSTER and MODAL, and have convinced us that interactive computer procedures should be established for improved cloud cover verification in future studies.

### 1. CONCLUSIONS

We have evaluated the 3DNEPH MODAL algorithm, along with an alternative CLI STER algorithm, and have found that CLI STER performs better. The CLUSTER algorithm finds clear/cloud boundaries, numbers of cloud layers, and fractions of cloud cover that are closer to those estimated by observers evaluating the IR and visible images. The MODAL algorithm finds more cloud layers than are observed, and tends to overanalyze regions of transition from one layer to another. The CLISTER algorithm is designed to assign all points to a layer in an optimal manner. The MODAL algorithm frequently (27 percent of the time) terminates its histogram analysis with unassigned points. These points are later assigned to warmer modes, which will tend to overestimate the altitudes of clouds in those modes (except cirriform categories), since the 3DNEPH uses the coldest gravshade in a mode to represent the entire mode. CLUSTER uses a larger array of satellite data, and appears to benefit from the extra information by improved detection of low clouds such as cumulus. CLUSTER, moreover, produces all the information necessary for cloud cover estimates at the same mesh size as the 3DNEPH. On the McIDAS, CLUSTER takes very nearly the same computer time as MODAL, so that its computational requirements are reasonable.

Despite the overall advantages of CLISTER, there is considerable room for improvement, and algorithm development leading to accurate cloud analysis should continue. In particular, CLISTER finds too many layers of cirrus clouds, and the fractional coverage of both cumulus and cirrus clouds usually differs from the apparent coverage on the satellite images. When thermal gradients are found in clear land or snow, it can also find two or three clusters when only one is desired. The MODAL analyses were slightly better in these areas.

Unlike MODAL, the CLUSTER algorithm can be tuned in a number of places to reduce sensitivity, and give results closer to those that actually exist. For instance, it might be possible to improve cirrus layering by increasing the interval sizes in the cold temperatures. Surface regions can be simplified; that is, analyzed into fewer clusters, by increasing the minimum size of a cluster. These are but two examples of how CLUSTER is open to tuning by making very simple alterations to the algorithm. There are other places in the algorithm code where simple modifications can be made to better model specific meteorological conditions that occur commonly on a global basis.

The VFGWC has recognized the limitations of MODAL, and the complexity in computer code, so that a substantially revised version has been written for the planned upgrade of the 3DNEPH known as the Real-Time Nephanalysis (RTNEPH).

With the continued cooperation of AFGWC, and drawing on our experience with MODAL and CLUSTER, we plan to evaluate the RTNEPH satellite analysis algorithm. We also plan to use the capabilities of the McDAS system for improving estimates of "cloud truth" for the evaluation and refinement of the cloud analysis algorithms. Finally, we shall propose the use of a modified CLUSTER algorithm, developed at AFGL, in replacement of, or in tandem with, the RTNEPH/3DNEPH image analysis algorithms.

# References

- 1. Coburn, A. R. (1971) Improved Three-Dimensional Nephanalysis, AFGWC Technical Memorandum 71-2.
- 2. Fye, F. K. (1978) The AFGWC Automated Cloud Analysis Model, AFGWC Technical Memorandum 78-002.
- Hawkins, Rupert S. (1980) A Clustering Technique for Satellite Image Analysis, Proc. 8th Conf. on Weather Forecasting and Analysis, Amer. Meteor. Soc., 115-118.
- 4. Hawkins, Rupert S. (1981) Objective Analysis of Satellite Cloud Imagery,
  Proc. 1981 International Geoscience and Remote Sensing Symposium,
  IEEE Geoscience and Remote Sensing Society, Volume 1, 477-482.

